

# Applying Bayesian Structural Inference to Pitch Sequences in Music

Anastasiya Salova

June 9, 2016

## Abstract

In this project, I am exploring the applications of Bayesian Structural Inference to pitch sequences in music. I provide some examples of applying the method to pitch sequences translated to a 8-letter alphabet. I also outline some possibilities of improving the ways to apply BSI to both generating music and inferring information quantities of given samples.

## 1 Music as a Complex System

The common elements of music are pitch which governs melody and harmony, rhythm, dynamics (loudness and softness), and the sonic qualities of timbre and texture. Different types of music can emphasize, de-emphasize, or omit certain elements. However, the interplay between these elements is always important. Language was one of the first systems that was studied using the notion of Shannon's entropy. Music has a lot of similarities with language. For instance, letters, or sounds for spoken language, are analogous to notes and their duration, words are analogous to measures and sequences, and sentences and phrases are analogous to musical phrases. Some of the tonal languages have even more similarities with music, because tone is a part of what conveys meaning within words. However, even in non-tonal languages intonation is important and has a wide range of functions such as indicating the attitudes and emotions of the speaker, signalling the difference between statements and questions, or focusing attention on particular grammatical aspects of a sentence. These similarities between music and languages appear to be very fundamental. For example, some studies suggest similarities in neural cortex activation caused by listening to musical phrases and phrases of spoken language as discussed, for instance, in [2].

These structural similarities are suggesting that similar methods can be used to study music and language. Studying the statistics of language and measuring its informational quantities has produced numerous results. For instance, in 1951, Shannon applied the concept of entropy to calculate the entropy of English letters. Since then, similar calculations have been performed to many other languages. Entropy of music has also been studied, for instance, in [4] and [5]. Entropy in music has also been studied with respect to human listeners' models of music in [6]. However, complexity in structure of music results in having to take into account various parameters and structure on different timescales, makes studying the informational theoretical properties of music nontrivial.

## 2 Information Measures in Music

Perception of music appears to be very fundamental to humans. Studies show that, for instance, even newborns can discriminate between major and minor chords in Western classical music [7]. The main

purpose of language is to convey meaning, however, for music, the purpose is more linked to surprise and enjoyability. Considering various informational quantities as applied to music is a way to quantify these concepts.

To do that, it is important to come up with a strategy to quantify information quantities in music. For language, there is an obvious way to define entropy, because letters can be used as the smallest building blocks. In music, the choice of alphabet is more ambiguous. As a result there are more options to define entropy. For instance, one can define the entropy of pitch sequences, or rhythmic entropy (which is expected to produce lower entropy estimates due to smaller alphabet size and less variation). Both of the characteristics provide useful information about music. To get started, I was applying BSI to calculate the entropy of pitch sequence in music.

### 3 Available Data

I was using music21 to access musical pieces. Music21 is a tool to computer-aided musicology that can serve a variety of purposes. In my project, I was using the scores already in the music21 library, however, potentially one can study music downloaded and converted to .mxl format. I picked Bach's chorales to focus on because their structure makes them both easy and interesting to study. In particular, it is appropriate to use Bach's chorales for studying pitch sequences, as most of the melody is contained in Soprano voice. From music21 files one can get several pieces of information: set of voices (like Soprano, Alto, etc.), pitches, note and pause durations, information about rests, chords, key signatures, etc. The information I extracted was the pitch sequence in the highest (Soprano) voice of each piece. The highest voice was chosen because most of the melody of the piece is contained there. I will use bwv190.7-inst to illustrate the type of data I was working with:

F5 F5 D5 E5 F5 G5 E5 E5 E5 D5 D5 C5 D5 E5 E5 E5 E5 F5 D5 D5 D5 C5 D5 E5 E5 D5 F5 F5 D5 E5 F5 G5 E5 E5 D5 D5 C5 D5 E5 E5 D5 C5 C5 C5 C5 D5 D5 F5 F5 D5 E5 F5 G5 F5 F5 E5 F5 D5 D5 E5 F5 F5 D5 E5 F5 G5 F5 F5 E5 F5 D5 D5 E5 F5 G5 F5 F5 E5 F5 G5 E5 E5 E5 D5 D5 C5

Because I was expecting to apply the BSI to all pieces independently from their key signature, I was converting the pieces from their pitch representation to their scale degree representation (see figure 1). In the scale degree representation, one can consider each pitch in relation to the tonic. For instance, in scale degree representation, the soprano voice of the above piece looks like:

['2', '2', '0', '1', '2', '7', '1', '1', '1', '0', '0', '6', '0', '1', '1', '1', '1', '2', '0', '0', '0', '6', '0', '1', '1', '0', '2', '2', '0', '1', '2', '7', '1', '1', '1', '1', '2', '0', '0', '0', '6', '0', '1', '1', '0', '6', '6', '6', '6', '6', '0', '0', '2', '2', '0', '1', '2', '7', '2', '2', '1', '2', '7', '7', '1', '2', '2', '0', '1', '2', '7', '1', '1', '1', '0', '0']

If a harmonic scale is used, it is reasonable to only use 7 symbols for the notes in the key signature and reserve the 8th for all the unexpected notes (corresponding to 7 in an above sequence). That won't be appropriate for many other genres, composers, or even some other Bach pieces. In a general case, it might be necessary to use 12 symbols, for example, when a chromatic scale is used. Using as few symbols as possible is preferable in case of limited computational time.

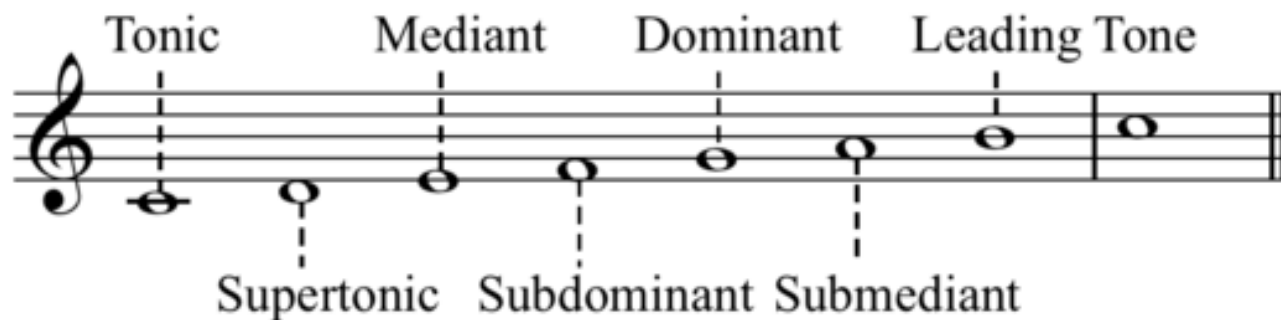


Figure 1: Scale degree

## 4 Bayesian Structural Inference Applied To Music

One of the most straightforward approaches to studying informational quantities is applying Markov model. In the simplest case one can assume that music is produced by a 1st order Markov process. That approach will provide conditional probabilities of observing or hearing each pitch given the previous one. That approach will provide some information about the piece. However, making a Markovity assumption is not a good idea in this case. For instance, that will completely overlook the connections between different sections of each piece, even though that structure is very important for music.

A slight improvement can be achieved by using n-grams consisting of pitch sequences which will provide a better approach to reconstructing musical sentences. It might also make sense to consider sequences of different length in one model. The beamed sequences of notes, for instance, can be considered as one element in a sequence. However, that will still not be a good way to pick up the structure of the piece on a larger scale.

One of the other ways to study music structure with a goal of obtaining information theoretical measures is Bayesian Structural Inference. Bayesian structural inference allows to infer hidden structure and organization from natural and designed systems. Music is a designed system that, however, heavily relies on human perception, which brings a natural system into play. That method doesn't assume Markovity in the input sequence, which is appropriate for pitch and duration sequences in music. Also, BSI provides various high-probability machines that could have generated the desired sequence, which is useful if one wants to find structural similarities in different music generators. Moreover, Bayesian Structural Inference is perfectly suited for generating musical sequences from the inferred  $\epsilon$ -machine.

BSI also contains a helpful tool in evaluating the method performance. Music is interesting due to the effect that it has on the listener, and listening to a generated sequence provides a chance to evaluate the output from that standpoint. There exist numerous algorithmic musical composition methods, but BSI provides access to the structure of the generator and requires relatively low computation time.

Finally, it is useful to apply Bayesian Structural Inference because various informational quantities and their distributions can be obtained over all the inferred machines given their probabilities. That, for instance, can allow to produce sequences of slightly different entropy measures that are based on the same piece.

## 5 Examples and Method Limitations

Bayesian Structural Inference Method is a way of inferring structure of the process from finite data series built with a finite state alphabet by means of inferring the  $\epsilon$ -machines that are most likely to have generated a given sequence. A very detailed introduction of the method can be found in the BSI for Hidden processes paper [3]. BSI is a two-step process, where the first step is finding the probability of each topology, and the second step is assigning transition probabilities within topological  $\epsilon$ -machines with that have nonzero probabilities to have generated a given sequence. BSI is easiest to apply to short alphabets when the resultant  $\epsilon$ -machine is expected to have a relatively low number of states.

Though the process of applying the BSI is not difficult conceptually given the CMPy tools, there are some computational time limits as mentioned above. Thus, I have not been able to apply the method to all Bach chorals available. Here, I present some examples of inferred *epsilon*-machines and sequences that have been produced applying BSI to a particular piece of music. I picked bwv190.7-inst as an example because its melody is pretty self-sufficient, it mostly contains the scale degree notes with a few exceptions, and the rhythmic structure is relatively simple. The last part is important for inspecting the generated piece. As I didn't change the rhythmic structure but changed the pitch sequence, getting a complicated rhythmic pattern is distracting.

As mentioned above, the number of topological  $\epsilon$ -machines grows exponentially in both the number of states and the number of symbols. There doesn't seem to be an easy straightforward way to evaluate the exact  $\epsilon$ -machine number, but it is clear that it is unreasonable to expect to be able to parse through all of them, even the number of machines is as low as 5.

The way to avoid that is to run the reconstruction algorithm on a random selection of several thousand machines. I tried to run the algorithm on 1000-5000 5-state machines. Most of the topological machines naturally are not compatible with the observed sequence. For instance, in case of bwv190.7-inst, only about 4% of the machines had a nonzero probability of producing a given sequence. After running the BSI algorithm, I got the list of machines that can possibly produce an observed sequence. To generate new pitch sequences, I only used the ones with the highest probability. The method is likely to demonstrate better results if more machines are compared, which is just a question of computation time assigned to the task. Here, I present the machines I got for bwv190.7-inst. Figure 2 corresponds to the reconstruction based on 1000 random samples, and figure 3 is based on 5000.

Then, I have made estimates of excess entropy and statistical complexity of the processes. The figures below demonstrate prior and posterior distributions for a 5000 random machine case (figure 4) as well as  $h_\mu$  vs  $C_\mu$  for both prior and posterior distributions (figure 5). The values of both  $C_\mu$  and  $h_\mu$  converge to approximately 2 bits in this case.

## 6 Further Work

In this section, I will outline some of the possible directions in which further investigate of applying BSI to music can go.

One of the possible BSI applications is setting limits on entropy measures for different genres of music. Maybe we can trace the historical progression of information theoretical quantities in music by genre and time period. However, it poses a challenge of evaluating the information theoretical quantities of genres arising from different structures.

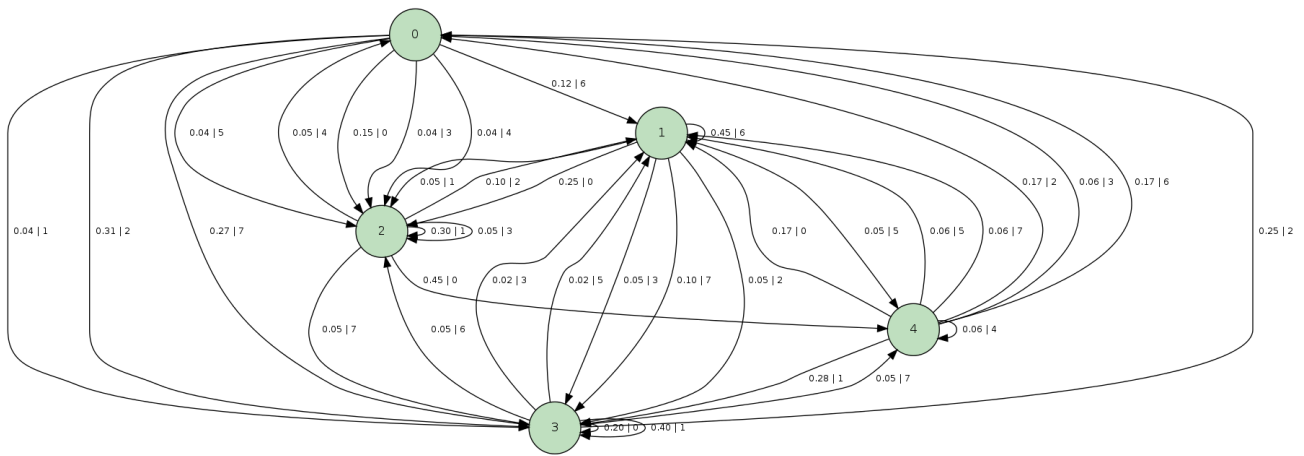


Figure 2: Best EM estimate from 1000 random samples

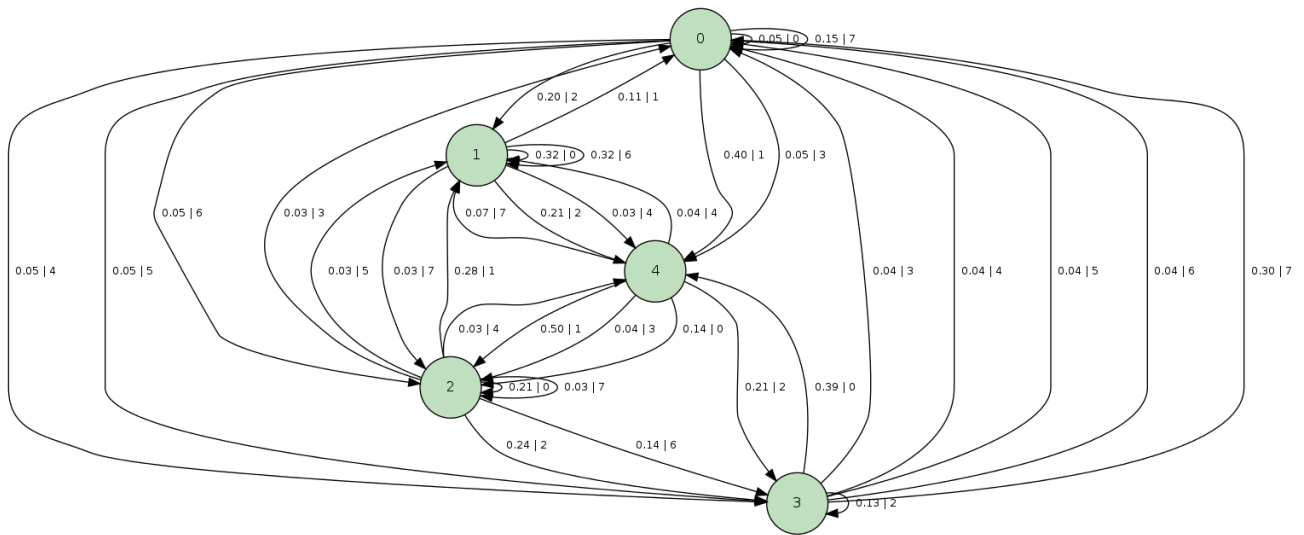


Figure 3: Best EM estimate from 5000 random samples

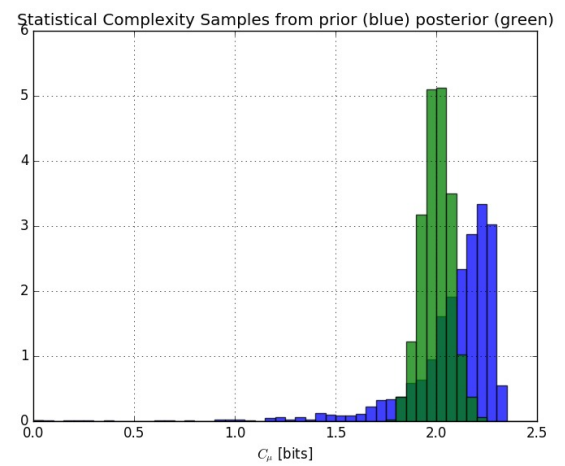
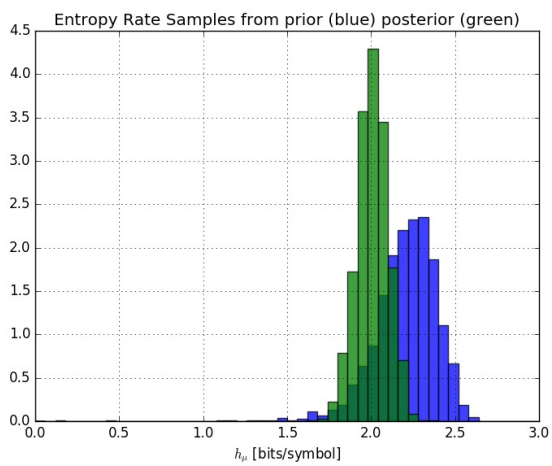


Figure 4:  $h_\mu$  (left) and  $C_\mu$  (right) for prior (blue) and posterior (green) distributions

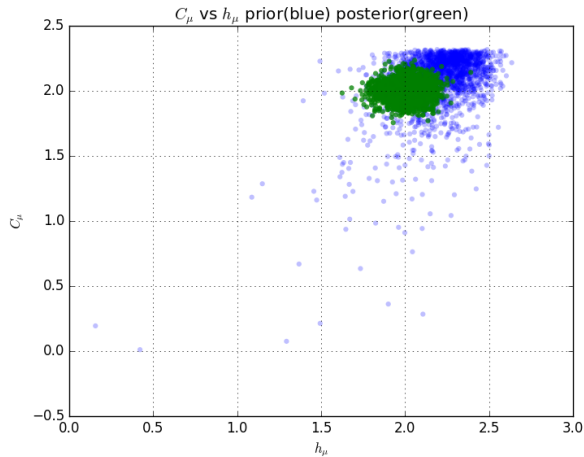


Figure 5:  $h_\mu$  vs  $C_\mu$  for prior (blue) and posterior (green) distribution

The other direction is studying how different types of entropy contribute to the structure of music. One can consider different types of entropy (e.g. rhythmic, pitch, chord progression) to be varying separately. However, because rhythmic structure and pitch structure of the piece are interrelated, there might be a way to come up with some joint entropy measure that will take that connection into account.

It will also be interesting to come up with ways to generate music using BSI in a more sophisticated way taking into account different parameters such as pitch and rhythm. The most straightforward way involves labelling the states using both the length of the note and the pitch. That way may not be optimal and involves a very large alphabet, which means very high computation time and reconstructing machines with a very large number of states. Another way involves considering the time series where the time step is the length of a shortest interval (pitch duration or rest duration) involved in the piece. Then the size of the alphabet doesn't increase and the training sequence becomes longer. The problem with applying that method is not differentiating between longer notes and a sequence of the same pitches that add up to the same length.

Finally, it would be interesting to apply BSI to separate music pieces and find the  $\epsilon$ -machines that best describe the entire set. That would be helpful, for instance, to see the correspondence between states or parts of the sequence generating  $\epsilon$ -machine and music theoretical rules. For instance, one could see if any part of the machine corresponds to cadence sequences.

## References

- [1] James P. Crutchfield, David P. Feldman, Regularities unseen, randomness observed: Levels of entropy convergence, *Chaos*, Volume 13, 2003.
- [2] S. Brown et al., Music and language side by side in the brain: a PET study of the generation of melodies and sentences, *European Journal of Neuroscience*, Volume 23, Issue 10, pages 2791–2803, May 2006
- [3] Christopher C. Strelhoff, James P. Crutchfield, Bayesian Structural Inference for Hidden Processes, *Physical Review E* 89 (2014) 042119
- [4] Gerardo Febres, Klaus Jaffé, Music Viewed by Its Entropy Content: A Novel Window for Comparative Analysis, [arxiv.org/abs/1510.01806](https://arxiv.org/abs/1510.01806)
- [5] Gregory E. Cox, On the Relationship Between Entropy and Meaning in Music: An Exploration with Recurrent Neural Networks, Conference: Proceedings of the 32nd Annual Conference of the Cognitive Science Society, August 2010
- [6] L. Manzara et al., On the Entropy of Music: An Experiment with Bach Chorale Melodies, *Leonardo Music Journal* Vol. 2, No. 1 (1992), pp. 81-88
- [7] P. Virtala et al., Newborn infants' auditory system is sensitive to Western music chord categories, *Front Psychol.* 2013; 4: 492